

The University of Chicago

THE MINIMUM OF A FUNCTION OF
INTEGRALS IN THE CALCULUS
OF VARIATIONS

A DISSERTATION SUBMITTED TO THE
FACULTY OF THE DIVISION OF THE
PHYSICAL SCIENCES IN CANDIDACY FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

1938

By

CARROLL PARKER BRADY

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

Reprinted from
CONTRIBUTIONS TO THE CALCULUS OF VARIATIONS, 1938-1941
The University of Chicago Press

1942

The University of Chicago

THE MINIMUM OF A FUNCTION OF
INTEGRALS IN THE CALCULUS
OF VARIATIONS

A DISSERTATION SUBMITTED TO THE
FACULTY OF THE DIVISION OF THE
PHYSICAL SCIENCES IN CANDIDACY FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

1938

By
CARROLL PARKER BRADY



Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

Reprinted from
CONTRIBUTIONS TO THE CALCULUS OF VARIATIONS, 1938-1941
The University of Chicago Press
1942

TABLE OF CONTENTS

Section	Page
Introduction	1
1. Formulation of the problem	2
2. The multiplier rule	5
3. The necessary conditions of Weierstrass and Clebsch	13
4. The second variation and the accessory minimum problem	15
5. The condition of Mayer	21
6. Sufficiency theorems from the theory of the problem of Bolza	28
7. Extension of the sufficiency proof	31
8. Application to a special problem	38
9. A second application	46
Bibliography	48

THE MINIMUM OF A FUNCTION OF INTEGRALS
IN THE CALCULUS OF VARIATIONS

Introduction. In the fourth chapter of his "Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes" Euler discusses a number of calculus of variations problems in which a function of several integrals is to be minimized or maximized [I, chap. 4].¹ For a general problem of that type [I, p. 161] he shows that the differential equations necessarily satisfied by a minimizing arc are the same as the equations that must be satisfied by an extremal arc for the isoperimetric problem of minimizing a single one of the integrals when the remaining integrals have prescribed values. Bolza [II] considered two of the special problems that Euler had discussed [I, pp. 148 and 153], and by means of the theories of associated isoperimetric problems he established the existence of arcs for which the functions of integrals considered took on strong relative extreme values. No doubt the method used by Bolza could be generalized so as to apply to the more general problem to be considered in this paper. The problems treated by Euler and Bolza may also be transformed into problems of Bolza in the calculus of variations. It is the purpose of this paper to apply the theory of the problem of Bolza to a general problem in which the functional to be minimized is a function of integrals. By so doing we develop a theory directly applicable to problems of the type discussed by Bolza and Euler.

¹The Roman numerals in square brackets refer to the bibliography at the end of this paper.

In Section 1 we formulate the problem to be considered in this paper, which we call problem A, and transform it into a problem B which is a problem of Bolza but not of the most general type. The method of the following pages is to interpret for problem A via the problem B the results already in the literature concerning the general problem of Bolza which is designated here as problem C. Section 2 below contains a discussion of the multiplier rule and of the extremals. We then consider, in Section 3, the conditions of Weierstrass and Clebsch. Preparatory to the formulation of two Mayer conditions in Section 5 we discuss the second variation and the accessory extremals in Section 4. An application of the sufficiency theorems for the problem of Bolza is made in Section 6. This procedure yields only a restricted strong relative minimum for problem A because the definitions of a strong relative minimum for problems A and B are not equivalent. A satisfactory theorem concerning strong relative minima for problem A can be deduced, however, by supplementing the results for problem B, as is done in Section 7, by methods similar to those used by Hestenes for isoperimetric problems. Finally, in Sections 8 and 9 we apply the more general theory here developed to the two special problems considered by both Euler and Bolza.

1. Formulation of the problem. As a generalization of the problems mentioned in the introduction and considered by Bolza we propose the following one. Consider a class $\mathcal{M}$ of so-called admissible arcs in the xy-plane defined by functions of the form

$$(1:0) \qquad y(x) \qquad (x_1 \leq x \leq x_2).$$

Let us suppose that for each of these arcs the integrals

$$(1:1) \qquad I_{\beta} = \int_{x_1}^{x_2} h_{\beta}(x, y, y') dx \qquad (\beta = 1, \dots, n)$$

and the function $g(I_1, \dots, I_n)$ have well defined values. Then the problem to be considered is that of finding in the sub-class of admissible arcs which join two fixed points in the xy -plane one which minimizes the function g . In the following pages this problem will be referred to as problem A.

In order to provide an analytic basis for the study of this problem we consider an arc E belonging to the above mentioned sub-class of $\mathcal{M}$ in which a minimizing arc is to be sought. We suppose that E is continuous and consists of a finite number of sub-arcs on each of which it has a continuously turning tangent. We assume that the functions $h_\beta(x, y, y')$ ($\beta = 1, \dots, n$) have continuous partial derivatives of at least the third order in a neighborhood R of the points (x, y, y') on E , that g has continuous partial derivatives of at least the second order in a neighborhood of the values $(I_1, \dots, I_n)$ belonging to E , and finally that at that point $(I_1, \dots, I_n)$ not all of the first derivatives of g vanish. An arc defined by functions of the form (1:0) will be said to be admissible if it is continuous and consists of a finite number of sub-arcs each of which has a continuously turning tangent, and if furthermore its elements (x, y, y') all lie interior to the region R . Such an arc will certainly give well defined values to the integrals I_β , and also to the function g provided that the region R is sufficiently small.

In the following pages we treat the foregoing problem A as a special case of the problem of Bolza in the calculus of variations. In order to do this we consider a class of arcs defined in an $(n + 2)$ -dimensional space by functions of the form

$$(1:2) \quad y_1(x) \quad (x_1 \leq x \leq x_2; i = 0, 1, \dots, n).$$

Such an arc will be termed admissible if it is continuous and

consists of a finite number of sub-arcs having continuously turning tangents, and if its xy_0 -projection, defined by the function $y_0(x)$, is admissible for the problem A described above. We designate the class of admissible arcs (1:2) by $\mathcal{M}'$. The problem of Bolza corresponding to the problem formulated in the first paragraph is then that of finding in the sub-class of the class $\mathcal{M}'$ satisfying the differential equations and end conditions

$$(1:3) \quad h_{\beta}(x, y_0, y_0') - y_{\beta}' = 0 \quad (\beta = 1, \dots, n),$$

$$(1:4) \quad x_s - a_s = y_0(x_s) - y_{0s} = y_{\beta}(x_1) = 0 \quad (s = 1, 2; \beta = 1, \dots, n)$$

one which minimizes the function

$$(1:5) \quad g[y_1(x_2), \dots, y_n(x_2)].$$

In the following pages this problem will be designated as problem B.

We note that the statements of problems A and B here given are entirely equivalent. This follows because there is a one-to-one correspondence between the arcs of the sub-class of $\mathcal{M}$ through the two given points 1 and 2 and the arcs of the sub-class of $\mathcal{M}'$ for which the conditions (1:3) and (1:4) are satisfied, and since furthermore on corresponding arcs the values of the function g are the same. We are justified therefore in applying to the more special problem A formulated above some of the well known results from the theory of the problem of Bolza.

In order to facilitate the statement of the results from the theory of the problem of Bolza that we desire to use, we here state that problem as it appears in the literature [VIII, pp. 9-11]. In the following pages it will be designated as problem C. The problem will be formulated in terms of the notation used in

the less general problem B just described. In stating problem B we were careful to make definitions and hypotheses that would ensure its appearance as an instance of the general problem of Bolza. We do not, therefore, give the similar though more extensive sets of definitions and hypotheses necessary to place the general theory on an analytic basis. With this understanding we then state the problem C of Bolza as that of finding in a class of admissible arcs defined in an $(n + 2)$ -space of points $(x, y_0, y_1, \dots, y_n)$ by functions of the form

$$(1:6) \quad y_i(x) \quad (x_1 \leq x \leq x_2; i = 0, 1, \dots, n)$$

and satisfying the differential equations and end conditions

$$(1:7) \quad \phi_\beta(x, y, y') = 0 \quad (\beta = 1, \dots, m < n + 1),$$

$$(1:8) \quad \psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \quad (\mu = 1, \dots, p \leq 2n + 4)$$

one which minimizes a sum

$$(1:9) \quad J = g[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} f[x, y(x), y'(x)] dx.$$

Here and elsewhere in the paper we use the notations $y, y(x)$ to represent the sets $(y_0, y_1, \dots, y_n), [y_0(x), y_1(x), \dots, y_n(x)]$ respectively, with similar notations for sets of derivatives. The equations (1:7) and (1:8) correspond to the equations (1:3), (1:4), with $m = n, p = n + 4$ in the less general case. The function g is a more general form of the corresponding function for problem B, and the integral containing f does not appear in problem B.

2. The multiplier rule. Among the conditions which must be satisfied along a minimizing arc for the general problem of Bolza the first to be studied is usually the multiplier rule.

That rule may be stated as follows [VIII, p. 26]:

An admissible solution E of the equations $\phi_\beta = 0$, defined on an interval $x_1 x_2$, is said to satisfy the multiplier rule if there exist constants l_0, e_μ not all zero and a function

$$(2:1) \quad F = l_0 f + l_\beta(x) \phi_\beta$$

with multipliers $l_\beta(x)$ continuous on $x_1 x_2$ except possibly at values x defining corners of E where they have well-defined right and left limits, such that equations of the form

$$(2:2) \quad F_{y_1'} = \int_{x_1}^x F_{y_1} dx + c_1$$

are satisfied along E , and furthermore such that the equation

$$(2:3) \quad [(F - y_1' F_{y_1'}) dx + F_{y_1'} dy_1]_{x_1}^{x_2} + l_0 dg + e_\mu d\psi_\mu = 0$$

holds at the ends of E for every choice of the differentials $dx_1, dy_{11}, dx_2, dy_{12}$. For an arc E satisfying the multiplier rule the multipliers $l_0, l_\beta(x)$ do not vanish simultaneously at any point of the interval $x_1 x_2$. Every minimizing arc E for the problem of Bolza must satisfy the multiplier rule.

In the above paragraph and hereafter it is understood that the ranges of the subscripts are $i = 0, 1, \dots, n$; $\beta = 1, \dots, m$; and $\mu = 1, \dots, p$, with $m = n$, $p = n + 4$ for the problem B. Terms with repeated indices will represent sums of such terms over a range of values associated with the given indices, as is customary in tensor analysis. The literal subscripts y_β, y_β' attached to F , or to other functions later on, indicate partial derivatives. Use is also made of the notation y_{11} to designate $y_1(x_1)$, and other similar notations.

In applying the multiplier rule to problem B we may first notice that the function F for that problem takes the form

$$(2:4) \quad F(x, y_0, y', \lambda) = \lambda_\beta(x) [h_\beta(x, y_0, y_0') - y_\beta'].$$

On imposing the condition that the function F shall satisfy the last n of the equations (2:2) along a minimizing arc E one finds that the multipliers have the values $\lambda_\beta(x) = c_\beta$. We shall accordingly hereafter designate the constant multipliers by λ_β .

If we define a function $h(x, y_0, y_0', \lambda)$ by the equation

$$(2:5) \quad h = \lambda_\beta h_\beta(x, y_0, y_0'),$$

then the first of the equations (2:2) may be written as

$$(2:6) \quad h_{y_0'} = \int_{x_1}^x h_{y_0} dx + c_0.$$

The condition that the coefficients of the various differentials in (2:3) vanish is called the transversality condition. For the problem B , in which F is the function (2:4) and ψ_μ the conditions (1:4), it may be seen that the transversality condition on a minimizing arc E is equivalent essentially to the n equations

$$(2:7) \quad -\lambda_0 + \lambda_0 g_\beta = 0,$$

in which the subscript β attached to g indicates the partial derivative of g with respect to I_β .

Due to the peculiar nature of the problem B we may show that if a minimizing arc E for that problem satisfies the multiplier rule then E is normal, that is, there is no non-null set of multipliers λ_0, λ_β having $\lambda_0 = 0$ with which E satisfies the rule. This follows from the equation (2:7), since if $\lambda_0 = 0$ then also $\lambda_\beta = 0$, or all the multipliers λ_0, λ_β vanish. We summarize the last two paragraphs in the following theorem:

THEOREM 2:1. Along a minimizing arc E for the problem A formulated in Section 1 there must exist a function $h(x, y_0, y_0', \lambda)$

of the form (2:5), with constant multipliers l_β , such that the equation (2:6) is satisfied identically along E. The multipliers l_0, l_β must further satisfy with E the transversality conditions (2:7) in which the arguments of g are $y_{\beta 2}$, or the values of the integrals I_β taken along the arc E. The arc E for which these conditions are satisfied is, moreover, normal, so that the multiplier l_0 can always be taken to be unity.

The multiplier rule has three important corollaries which will be found useful in the succeeding discussion. Due to the special properties of our particular problem B the results stated in these corollaries appear here in a somewhat simpler form than they do in the theory of the more general problem C of Bolza [VIII, p. 27.1]. The first corollary contains the statement that at each point of an admissible arc E satisfying the equation (2:6) the function h_{y_0} has forward and backward derivatives, equal except possibly at corners, and such that the following Euler-Lagrange equation holds:

$$(2:8) \quad dh_{y_0} / dx = h_{y_0}.$$

The second corollary provides us with the Weierstrass-Erdmann corner condition, which states that at each corner of an admissible arc E satisfying the equation (2:6) the function h_{y_0} has well-defined right and left limits which are equal. The third corollary, whose result is known as the Hilbert differentiability condition, implies that the functions $y_1'(x)$ defining an admissible arc E have continuous derivatives of at least the first order with respect to x on each sub-arc of E satisfying the equations (2:6) and $h_\beta - y_\beta' = 0$ on which the functions $y_1(x)$ defining E have continuous first derivatives and the function h_{y_0}, y_0' is different from zero. The first two corollaries are immediate

consequences of the equation (2:6). The proof of the third corollary involves an application of implicit function theory.

We have seen that a minimizing arc for problem B not only satisfies the equations $h_\beta - y_\beta' = 0$ but also must have associated with it constant multipliers λ_β with which it satisfies the equation (2:8). Solutions of these equations play an important role in the development of the theory. Bliss [VIII, pp. 33-39], and Bliss and Hestenes [V, pp. 309-311], have made studies of the solutions of the analogous sets of differential equations occurring in the more general theories of the problems of Bolza and Mayer, both of which apply to the problem B. However, there are some relations peculiar to problem B that make possible specialized results which we find it desirable to present. In order to study this phase of the problem we define an extremal to be an admissible arc $y_0(x)$, $y_\beta(x)$ and a set of constant multipliers λ_β , having continuous second derivatives and satisfying the equations

$$h_\beta - y_\beta' = 0, \quad dh_{y_0}/dx - h_{y_0} = 0.$$

A non-singular extremal is one along which the function h_{y_0}, y_0' is different from zero. In order to discuss such extremals we introduce so-called canonical variables (x, y, z) related to the variables (x, y, y', λ) by the equations

$$(2:9) \quad \begin{aligned} z_0 &= F_{y_0'} = h_{y_0'}(x, y_0, y_0', \lambda), \\ z_\beta &= F_{y_\beta'} = -\lambda_\beta, \quad 0 = h_\beta - y_\beta' \end{aligned}$$

whose determinant with respect to the variables (y', λ) has, except for a non-vanishing constant factor, the value $h_{y_0'} y_\beta'$. In a neighborhood of the elements (x, y, z) belonging to a non-singular extremal E the equations (2:9) have then a solution of the form

$$(2:10) \quad \begin{aligned} y_0' &= P(x, y_0, z), & y_\beta' &= f_\beta(x, y_0, P), \\ \ell_\beta &= -z_\beta = L_\beta(z_\beta), \end{aligned}$$

reducing to $y_1'(x)$, ℓ_β at each point $[x, y(x), z(x)]$ belonging to E . The function $P(x, y_0, z)$ is found by solving the first equation in (2:9) for y_0' and then substituting $-z_\beta$ for ℓ_β in the solution. The differential equations satisfied by the functions $y_1(x)$, $z_1(x)$ defining an extremal are

$$(2:11) \quad \begin{aligned} y_0' &= P(x, y_0, z), & z_0' &= h_{y_0}(x, y_0, P, L), \\ y_\beta' &= h_\beta(x, y_0, P), & z_\beta' &= 0. \end{aligned}$$

These equations follow directly from the equations (2:10), the definition of the z 's, and the Euler-Lagrange equation (2:8). The above equations have an interesting and useful homogeneity property. Since h is homogeneous of degree one in the ℓ 's and the solutions (2:10) of equations (2:9) are unique, it follows that for an arbitrary multiplier k

$$P(x, y, kz) = P(x, y, z).$$

Hence if the functions $y_1(x)$, $z_1(x)$ are the solution of equations (2:11) uniquely determined by the initial values a_1 , b_1 at $x = x_1$, then the functions $y_1(x)$, $kz_1(x)$ will be the solutions corresponding to the initial values a_1 , kb_1 .

In the theory of the general problem C the extremals are discussed with the help of the canonical variables (x, y, z) defined by the equations

$$(2:12) \quad z_1 = F_{y_1}, \quad \phi_\beta = 0.$$

These equations have solutions in a neighborhood of a non-singular extremal of the form

$$(2:13) \quad y_1' = P_1(x, y, z), \quad \ell_\beta = L_\beta(x, y, z).$$

The differential equations satisfied by the functions $y_1(x)$, $z_1(x)$ defining an extremal are then found to be the equations

$$(2:14) \quad y_1' = P_1(x, y, z), \quad z_1' = F_{y_1}[x, y, P(x, y, z), L(x, y, z)].$$

By means of these equations it is proved in the general theory [VIII, pp. 36-37] that every non-singular extremal arc E for a fixed value of the multiplier ℓ_0 is imbedded for values

$$(2:15) \quad x_1 \leq x \leq x_2, \quad a_{i0}, \quad b_{i0} \quad (i = 0, 1, \dots, n)$$

in a $(2n + 2)$ -parameter family of extremals defined by functions of the form

$$(2:16) \quad y_1(x, a, b), \quad z_1(x, a, b)$$

in which a, b are symbols for $(n + 1)$ -dimensional sets $(a_0, a_1, \dots, a_n)$, $(b_0, b_1, \dots, b_n)$. The functions y_1, y_{1x}, z_1, z_{1x} have continuous partial derivatives of at least the second order in a neighborhood of the values (x, a, b) defining E , and the determinant

$$(2:17) \quad \begin{vmatrix} y_{1a_k} & y_{1b_k} \\ z_{1a_k} & z_{1b_k} \end{vmatrix}$$

is different from zero along E . In terms of the variables (x, y, y', ℓ) the extremals (2:16) are defined by functions

$$(2:18) \quad y_1(x, a, b), \quad \ell_\beta(x, a, b)$$

related to the functions (2:16) by the equations (2:12) and such that the functions y_1, y_{1x}, ℓ_β have continuous partial derivatives of at least the second order.

The analogous results for problem B are contained in the following theorem:

THEOREM 2:2. For the problem B, which is equivalent to the problem A, the equations (2:9), (2:10) and (2:11) correspond respectively to the equations (2:12), (2:13) and (2:14) belonging to the general problem C. The functions (2:16) defining a $(2n + 2)$ -parameter family of extremals in which a given non-singular extremal E is imbedded may be written for the problem B in the form

$$(2:19) \quad \begin{aligned} & y_0(x, a_0, b_0, \ell), & z_0(x, a_0, b_0, \ell), \\ & y_\beta(x, a_0, b_0, \ell) + a_\beta, & z_\beta = -\ell_\beta, \end{aligned}$$

in which the functions y_0, z_0 are solutions of the first two equations in (2:11) with the variables z_β regarded as constant parameters, and in which

$$y_\beta(x, a_0, b_0, \ell) = \int_{x_1}^x h_\beta(x, y_0, P) dx + a_\beta.$$

The functions y_1 have the homogeneity property

$$y_1(x, a_0, kb_0, k\ell) = y_1(x, a_0, b_0, \ell).$$

The determinant (2:17) has the form

$$(2:20) \quad \begin{vmatrix} y_{0a_0} & 0 & y_{0b_0} & y_{0\ell_\gamma} \\ y_{\beta a_0} & \delta_{\beta\gamma} & y_{\beta b_0} & y_{\beta\ell_\gamma} \\ z_{0a_0} & 0 & z_{0b_0} & z_{0\ell_\gamma} \\ 0 & 0 & 0 & -\delta_{\beta\gamma} \end{vmatrix}$$

$$(\gamma, \beta = 1, \dots, n; \quad \delta_{\beta\gamma} = 0 \text{ if } \gamma \neq \beta; \quad \delta_{\beta\beta} = 1).$$

This determinant is different from zero along E. In terms of the variables (x, y, y', ℓ) the imbedding family (2:19) is defined by

$$(2:21) \quad y_0(x, a_0, b_0, \ell), \quad y_\beta(x, a_0, b_0, \ell) + a_\beta.$$

On account of their homogeneity property the functions
 $y_1(x, a_0, b_0, \ell)$ depend only upon the parameter a_0 and the n
ratios of the parameters b_0, ℓ_β . The functions appearing in
 (2:19) and (2:21) have continuity properties like those of the
corresponding functions (2:16) and (2:18).

3. The necessary conditions of Weierstrass and Clebsch.

Two further conditions playing important roles in the theory of the problem of Bolza are the necessary conditions of Weierstrass and Clebsch. These conditions may be stated for problem C as follows [VIII, pp. 63 and 65]:

An admissible arc E satisfying the equations $\phi_\beta = 0$ and the multiplier rule, with multipliers $\ell_0 = 1$, $\ell_\beta(x)$, is said to satisfy the necessary condition of Weierstrass with these multipliers if the condition

$$(3:1) \quad E(x, y, y', \ell, Y') \geq 0,$$

where

$$E = F(x, y, Y', \ell) - F(x, y, y', \ell) - (Y_1' - y_1')F_{y_1'}(x, y, y', \ell),$$

is valid at every element (x, y, y', ℓ) of E for all admissible sets $(x, y, Y') \neq (x, y, y')$ satisfying the equations $\phi_\beta = 0$.

The arc E is further said to satisfy the necessary condition of Clebsch with the multipliers $\ell_0 = 1$, $\ell_\beta(x)$ if the condition

$$(3:2) \quad F_{y_1' y_k'}(x, y, y', \ell) \pi_1 \pi_k \geq 0 \quad (1, k = 0, 1, \dots, n)$$

is valid at every element (x, y, y', ℓ) of E for all sets

$(\pi_0, \dots, \pi_n) \neq (0, \dots, 0)$ satisfying the equations

$$(3:3) \quad \phi_{\beta y_1'}(x, y, y', \ell) \pi_1 = 0.$$

Every normal minimizing arc for the problem of Bolza must satisfy these two conditions.

For the problem B the function E appearing in (3:1) takes the form

$$(3:4) \quad E(x, y_0, y_0', Y_0', \ell) = h(x, y_0, Y_0', \ell) - h(x, y_0, y_0', \ell) \\ - (Y_0' - y_0')h_{y_0'}(x, y_0, y_0', \ell),$$

as one readily verifies by use of the equations (2:4) and (2:5). The function E is here free from the arguments $y_\beta, y_\beta', Y_\beta'$. If it vanishes only for sets (x, y_0, Y_0') which are equal to (x, y_0, y_0') , it will also vanish only for full sets (x, y, Y') which are equal to (x, y, y') . This follows from the fact that the equations $\phi_\beta = 0$ for the problem B are the equations $h_\beta - y_\beta' = 0$, and hence the set (x, y, Y') is different from the set (x, y, y') if and only if the set (x, y_0, Y_0') is different from the set (x, y_0, y_0') . We see then that the Weierstrass condition for the problem B may be expressed in terms of the (xy_0) -projection of the minimizing arc E. A similar remark holds also for the Clebsch condition since for problem B the equations (3:3) are

$$h_{y_0'}(x, y_0, y_0', \ell)\pi_0 - \pi_\beta = 0 \quad (\beta = 1, \dots, n).$$

Therefore the condition $(\pi_0, \dots, \pi_n) \neq (0, \dots, 0)$ implies $\pi_0 \neq 0$. By the use of equations (2:4) and (2:5) one finds that for problem B the inequality (3:2) is

$$h_{y_0' y_0'}(x, y_0, y_0', \ell)\pi_0^2 \geq 0.$$

Since $\pi_0 \neq 0$, the Clebsch condition implies that along every minimizing arc for problem B that satisfies the multiplier rule the inequality $h_{y_0' y_0'} \geq 0$ must be satisfied. The word normal appearing in the statement of the Weierstrass and Clebsch

conditions may be deleted in the case of problem B, since for that problem arcs satisfying the multiplier rule are all normal. The results of this section are summarized for the problem A in the following theorem:

THEOREM 3:1. Every minimizing arc E for the problem A, where E is an admissible arc satisfying the multiplier rule with multipliers $\lambda_0 = 1, \lambda_\beta$, must satisfy with these multipliers the inequality

$$E(x, y_0, y_0', Y_0', \ell) \geq 0$$

at every element (x, y_0, y_0', ℓ) of E for all admissible sets $(x, y_0, Y_0') \neq (x, y_0, y_0')$. The function $E(x, y_0, y_0', Y_0', \ell)$ is the function defined by the equation (3:4). Further, along the minimizing arc E the inequality

$$(3:5) \quad h_{y_0', y_0'}(x, y_0, y_0', \ell) \geq 0$$

must be satisfied.

4. The second variation and the accessory minimum problem.

Another condition which must necessarily be satisfied by a minimizing arc for the problem of Bolza is the condition of Mayer, two forms of which will be described in the next section. The condition of Mayer appears in the literature as a criterion for ensuring the non-negativeness of the so-called second variation. In order to discuss the second variation, we first define a set of admissible variations as a set of functions $\eta_1(x)$ which have derivatives and continuity properties like those of the functions $y_1(x)$. For the problem C the second variation $J_2(\xi, \eta)$ of J along E may then be defined to be

$$(4:1) \quad J_2 = 2\gamma[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] + \int_{x_1}^{x_2} \omega(x, \eta, \eta') dx,$$

where ξ_1, ξ_2 are constants. The term 2γ outside the integral sign is found by substituting

$$dx_s = \xi_s, \quad dy_{1s} = y_1'(x_s)\xi_s + \eta_1(x_s) \quad (s = 1, 2)$$

in the quadratic form

$$2\gamma = [(F - y_1'F_{y_1})dx + 2F_{y_1}dy_1]_1^2 + 2q + 2e_\mu q_\mu$$

in which the constants e_μ are those of the multiplier rule and $2q$ and $2q_\mu$ are the quadratic forms in dx_s, dy_{1s} whose coefficients are the second derivatives of g and ψ_μ , respectively. The integrand function $2\omega(x, \eta, \eta')$ is the quadratic form

$$2\omega = F_{y_1 y_k} \eta_1 \eta_k + 2F_{y_1 y_k'} \eta_1 \eta_k' + F_{y_1' y_k'} \eta_1' \eta_k' \\ (i, k = 0, 1, \dots, n).$$

The equations

$$(4:2) \quad \Phi_\beta(x, \eta, \eta') = \phi_{\beta y_1} \eta_1 + \phi_{\beta y_1'} \eta_1' = 0,$$

$$(4:3) \quad \Psi_\mu[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] \\ = \psi_{\mu x_s} \xi_s + \psi_{\mu y_{1s}} [y_1'(x_s) \xi_s + \eta_1(x_s)] = 0$$

are called the equations of variation. From the theory of the problem C we now quote the following theorem [VIII, p. 72]:

For every normal non-singular minimizing arc E without corners the second variation $J_2(\xi, \eta)$ in equation (4:1) must be positive or zero for all admissible sets of variations $\eta_1(x)$ and constants ξ_1, ξ_2 which satisfy the equations of variation

$$\Phi_\beta = \Psi_\mu = 0 \text{ along } E.$$

For the problem B the equations of variation and the second variation take on a somewhat simpler appearance. If the functions $h_\beta - y_\beta'$ are substituted for the functions ϕ_β in the equations (4:2), we find the equations

$$(4:4) \quad h_{\beta y_0} \eta_0 + h_{\beta y_0'} \eta_0' - \eta_{\beta}' = 0 \quad (\beta = 1, \dots, n).$$

By use of the equations (1:4) it may be seen that the equations (4:3) reduce for problem B to the $n + 4$ equations

$$(4:5) \quad \xi_s = \eta_0(x_s) = \eta_{\beta}(x_1) = 0 \quad (s = 1, 2; \beta = 1, \dots, n).$$

By making use of the specialized functions (1:5) and (2:5) and the linearity of the end conditions ψ_{μ} in (1:4), plus the fact that $\xi_1 = \xi_2 = 0$, one finds that the quadratic forms 2γ and 2ω are, for the problem B,

$$(4:6) \quad 2\gamma[\eta(x_2)] = g_{\beta\gamma} \eta_{\beta}(x_2) \eta_{\gamma}(x_2) \quad (\gamma, \beta = 1, \dots, n),$$

$$(4:7) \quad 2\omega[x, \eta_0, \eta_0'] = h_{y_0 y_0} \eta_0^2 + 2h_{y_0 y_0'} \eta_0 \eta_0' + h_{y_0' y_0'} \eta_0'^2.$$

These remarks we extend in the following theorem:

THEOREM 4:1. For every non-singular minimizing arc E without corners the second variation for the problem B or A is

$$(4:8) \quad J_2(\eta) = g_{\beta\gamma} \eta_{\beta}(x_2) \eta_{\gamma}(x_2) + \int_{x_1}^{x_2} [h_{y_0 y_0} \eta_0^2 + 2h_{y_0 y_0'} \eta_0 \eta_0' + h_{y_0' y_0'} \eta_0'^2] dx$$

where

$$(4:9) \quad \eta_{\beta}(x) = \int_{x_1}^x (h_{\beta y_0} \eta_0 + h_{\beta y_0'} \eta_0') dx \quad (\beta = 1, \dots, n).$$

This variation must be positive or zero for every admissible variation $\eta_0(x)$ satisfying the equations $\eta_0(x_1) = \eta_0(x_2) = 0$.

This result is an easy consequence of the application to problem B of the corresponding theorem described above for problem C.

In the theory of the problem C the fact that the second variation must be positive or zero leads to an auxiliary minimum problem called the accessory minimum problem. This is the problem

of finding in the class of admissible sets of variations $\xi_1, \xi_2, \eta_1(x)$ satisfying the equations of variation $\Phi_\beta = 0, \Psi_\mu = 0$ along E , one which minimizes the second variation $J_2(\xi, \eta)$. The differential equations associated with this problem are

$$(4:10) \quad d\Omega_{\eta_1}/dx - \Omega_{\eta_1} = 0, \quad \Phi_\beta = 0,$$

where Ω is the homogeneous quadratic form in the variables $\eta_1, \eta_1', \lambda_\beta$ defined by the equation

$$(4:11) \quad \Omega(x, \eta, \eta', \lambda) = \omega + \lambda_\beta \Phi_\beta.$$

These equations are analogous to the equations $\phi_\beta = 0$ and $dF_{y_1}/dx = F_{y_1}$, the latter set of equations being derived from the equations (2:2) by differentiation of both members in each equation. The equations (4:10) are called the accessory differential equations. The transversality condition for the accessory minimum problem, analogous to the condition (2:3), is known as the accessory transversality condition.

By an accessory extremal is meant a solution $\eta_1(x), \lambda_\beta(x)$ ($x_1 \leq x \leq x_2$) of equations (4:10) having continuous derivatives $\eta_1', \eta_1'', \lambda_\beta'$ on $x_1 x_2$. A study of these extremals has been made by Bliss [VIII, pp. 73-76]. He introduces canonical variables (x, η_1, ξ_1) related to the variables $(x, \eta_1, \eta_1', \lambda_\beta)$ by means of the equations

$$(4:12) \quad \xi_1 = \Omega_{\eta_1'}(x, \eta, \eta', \lambda), \quad \Phi_\beta(x, \eta, \eta') = 0.$$

Bliss has shown that if E is a non-singular extremal then the equations (4:12) with $\lambda_0 = 1$ can be solved for the variables η_1', λ_β . These equations are of the form

$$(4:13) \quad \eta_1' = \Pi_1(x, \eta, \xi), \quad \lambda_\beta = \Lambda_\beta(x, \eta, \xi).$$

The equations (4:10) with $\lambda_0 = 1$ are then found to be equivalent to canonical accessory equations of the form

$$(4:14) \quad \eta_1' = \mathcal{H}_{\xi_1}(x, \eta, \xi), \quad \xi_1' = -\mathcal{H}_{\eta_1}(x, \eta, \xi),$$

in which $\mathcal{H}(x, \eta, \xi)$ is the homogeneous quadratic function in the variables η, ξ defined by the formula

$$(4:15) \quad \mathcal{H}(x, \eta, \xi) = \xi_1 \pi_1 - \Omega(x, \eta, \pi, \Lambda).$$

A fundamental set of $2n + 2$ linearly independent solutions of equations (4:14) is given by the elements of the columns of the determinant (2:17), with $\eta_1 = y_{1a_k}$, $\xi_1 = z_{1a_k}$ [VIII, p. 76].

For the problem B the function Ω has the form

$$(4:16) \quad \Omega(x, \eta_0, \eta', \lambda) = \omega(x, \eta_0, \eta_0') + \lambda_\beta [h_{\beta y_0} \eta_0 + h_{\beta y_0'} \eta_0' - \eta_\beta'],$$

where ω is the function defined in (4:7). The multipliers λ_β occurring here are constants. This follows from the fact that Ω does not contain η_β , so that the accessory equations (4:10) yield $d(\Omega_{\eta_\beta})/dx = d(-\lambda_\beta)/dx = 0$ for the problem B. If in the equations (2:3) we replace the function F by the function Ω defined in (4:16), the function g by the quadratic form 2γ defined in (4:6), and the functions ψ_μ by the functions $\bar{\Psi}_\mu$ defined in (4:5), we find that the accessory transversality condition for the problem B is equivalent essentially to the n equations

$$(4:17) \quad \xi_\beta + g_{\beta\gamma} \eta_\gamma(x_2) = 0 \quad (\beta, \gamma = 1, \dots, n),$$

where the $\xi_\beta = \Omega_{\eta_\beta} = -\lambda_\beta$ are constants along a given accessory extremal. According to Theorem 2:2 the determinant (2:20) is for problem B the determinant (2:17) of problem C. The elements of the columns of (2:20) therefore form a fundamental set of solutions

η_1, ζ_1 of the canonical accessory equations (4:14) calculated for problem B. These equations will be useful in the next section. To develop them, we notice that for problem B the equations (4:12) are

$$(4:18) \quad \begin{aligned} \zeta_0 &= h_{y_0 y_0'} \eta_0 + h_{y_0' y_0'} \eta_0' + \lambda_\beta h_{\beta y_0'}, & \zeta_\beta &= -\lambda_\beta \\ h_{\beta y_0} \eta_0 + h_{\beta y_0'} \eta_0' - \eta_\beta' &= 0, \end{aligned}$$

so that in the solutions (4:13) the functions π_1, Λ_β have the values

$$(4:19) \quad \begin{aligned} \pi_0(x, \eta_0, \zeta) &= [\zeta_0 - h_{y_0 y_0'} \eta_0 + \zeta_\beta h_{\beta y_0'}] / h_{y_0' y_0'}, \\ \pi_\beta(x, \eta_0, \zeta) &= h_{\beta y_0} \eta_0 + h_{\beta y_0'} \pi_0, & \Lambda_\beta &= -\zeta_\beta. \end{aligned}$$

The function $\mathcal{H}$ defined in (4:15) is then for problem B

$$(4:20) \quad \mathcal{H}(x, \eta_0, \zeta) = \zeta_0 \pi_0 + \zeta_\beta [h_{\beta y_0} \eta_0 + h_{\beta y_0'} \pi_0] - \omega(x, \eta_0, \pi_0),$$

in which the function ω has the definition (4:7). With this value for the function $\mathcal{H}$ the equations (4:14) become, in detail,

$$(4:21) \quad \begin{aligned} \eta_0' &= \pi_0(x, \eta_0, \zeta), & \zeta_0' &= -\zeta_\beta h_{\beta y_0} + h_{y_0 y_0'} \eta_0 + h_{y_0 y_0'} \pi_0, \\ \eta_\beta' &= h_{\beta y_0} \eta_0 + h_{\beta y_0'} \pi_0, & \zeta_\beta' &= 0. \end{aligned}$$

One readily verifies that one solution of the equations (4:21) is

$$(4:22) \quad \eta_1 \equiv 0, \quad \zeta_0 = h_{y_0'}, \quad \zeta_\beta = -\lambda_\beta,$$

where the constants λ_β are the multipliers belonging to E. We summarize the foregoing statements in the following theorem:

THEOREM 4:2. The canonical accessory equations for the problem B, which is equivalent to problem A, are the equations (4:21). One may solve the first two of these equations for η_0, ζ_0 as functions of x and the parameters ζ_β , from which it follows that the solutions η_β are of the form

$$(4:23) \quad \eta_{\beta}(x) = \int_{x_1}^x [h_{\beta y_0} \eta_0 + h_{\beta y_0'} \eta_0'] dx + k_{\beta},$$

where the k_{β} are constants. The columns of the determinant (2:20) form a complete set of solutions of the equations (4:21); one solution in particular is given in (4:22). The accessory transversality conditions for the problem B, or A, are given by the equations (4:17), in which the functions η_{β} and the constants $\zeta_{\beta} = -\lambda_{\beta}$ are solutions of the equations (4:21).

5. The condition of Mayer. In the literature there appear four closely related forms of the so-called condition of Mayer for the problem of Bolza in the calculus of variations. These are all conditions necessary for the non-negativeness, and in a strengthened form sufficient for the positiveness, of the second variation, as stated for problem B in Theorem 4:1. Of these four forms we shall make use of but one, that one being due to Bliss. For the less general problem of Bolza in which one end point of the admissible arcs is fixed Hestenes has defined a focal point condition which is also a condition of Mayer. Since the criterion for a focal point is comparatively simple theoretically, and since the problem B is a problem of Bolza of the type for which a focal point condition is defined, we also discuss that condition in this section.

First, we shall have need of a few definition. By the order q of abnormality of E on an interval $x_1 x_2$ is meant the number q of linearly independent sets of multipliers of the form $\ell_0 = 0, \ell_{\beta}(x)$ with which E satisfies the conditions (1:7) and (2:2) on $x_1 x_2$. We next define a function $Q(z)$ in terms of the second variation. After the usual integration by parts the second variation $J_2(\xi, \eta)$ along an accessory extremal is expressible as a quadratic form in ξ_1, ξ_2 and the end values of the

functions $\eta_1(x)$, $\zeta_1(x)$ defining the extremal. Let $\xi_{1\tau}$, $\xi_{2\tau}$, $\eta_{1\tau}$, $\zeta_{1\tau}$ ($\tau = 1, \dots, t$) be a maximum set of linearly independent constants and accessory extremals containing only sets ξ_1 , ξ_2 , η_1 , ζ_1 which are linearly independent of the special ones for which $\xi_1 = 0$, $\xi_2 = 0$, $\eta_1(x) \equiv 0$, and which satisfy the equations $\Psi_\mu = 0$. For the set

$$(5:1) \quad \xi_1 = \xi_{1\tau} z_\tau, \quad \xi_2 = \xi_{2\tau} z_\tau, \quad \eta_1 = \eta_{1\tau} z_\tau, \quad \zeta_1 = \zeta_{1\tau} z_\tau$$

the second variation may then be expressed as the quadratic form

$$(5:2) \quad Q(z) = J_2(\xi_\tau z_\tau, \eta_\tau z_\tau)$$

in the constants ($z_1, \dots, z_t$). A value $x_3 \neq x_1$ is said to define a point 3 conjugate to 1 if there exists an accessory extremal η_1 , ζ_1 satisfying the equations $\eta_1(x_1) = 0$, having $\eta_1(x_3) = 0$, and having the set $(\eta) \not\equiv 0$ on $x_1 x_3$. It is known [VII, p. 804, Theorem 5:1] that points 3 conjugate to 1 are the points other than 1 at which the matrix

$$(5:3) \quad \begin{vmatrix} \eta_{1j}(x_3) \\ \eta_{1j}(x_1) \end{vmatrix}$$

has rank less than $2n + 2 - q$, where η_{1j} forms with ζ_{1j} ($i = 0, 1, \dots, n$; $j = 1, \dots, 2n + 2$) a complete set of solutions of the canonical accessory equations, and q is the number of linearly independent sets η_1 , ζ_1 having $\eta_1(x) \equiv 0$ on $x_1 x_3$. This number q is also the order of abnormality of E on the interval $x_1 x_3$.

We now quote from the theory for problem C [VII, p. 802, Theorem 4:3] the result that if the extremal E is non-singular and the second variation J_2 is non-negative along E , then the condition $Q(z) \geq 0$ must hold for all sets of constants z not all zero;

moreover, if the order of abnormality of E is the same on every sub-interval x_3x_2 of x_1x_2 , then there can be no point 3 conjugate to 1 on E between its end points 1 and 2. This condition may be called the condition of Mayer as formulated by Bliss for the general problem C of Bolza.

If for problem C stated in Section 1 the first $(n + 1)$ of the equations $\psi_\mu = 0$ fix the coordinates of the initial points of admissible arcs at x_1 and the remaining $(p - n - 1)$ of the equations are conditions only on the coordinates of the end points of the arcs at x_2 , then the problem is called a problem of Bolza with second end point variable. For this problem a value $x_3 \neq x_2$ is said to define a focal point 3 of the manifold $\psi_\nu = 0$ ($\nu = n + 2, \dots, p$) on the extremal arc E if there exists an accessory extremal η_1, ξ_1 satisfying with a constant ξ_2 the equations

$$(5:4) \quad \bar{\Psi}_\nu = 0 \quad (\nu = n + 2, \dots, p)$$

and the accessory transversality condition at x_2 , having $\eta_1(x_3) = 0$, and having the set $(\eta) \neq (0)$ on x_3x_2 . From the theory of problem C [X, p. 545] we know that on a non-singular extremal arc E for which the second variation J_2 is non-negative and for which the abnormality of E is the same on every sub-interval of x_1x_2 , there can be no focal point 3 of the manifold $\psi_\nu = 0$ ($\nu = n + 1, \dots, p$) between the end points 1 and 2 of E . This may be called the focal point condition for a problem of Bolza with second end point variable.

In order to apply these results to problem B we first note that this problem is one with only the second end point variable. The equations (5:4) for problem B are simply

$$(5:5) \quad \xi_2 = 0, \quad \eta_0(x_2) = 0,$$

as we see from equations (4:5). To calculate $Q(z)$ as defined in (5:2) we note that for an accessory extremal η_1, ζ_1 satisfying the conditions (4:5), and with the function ω defined in (4:7), one finds

$$\int_{x_1}^{x_2} 2\omega dx = \int_{x_1}^{x_2} 2\Omega dx = \eta_1 \zeta_1|_1^2 = \eta_\beta(x_2) \zeta_\beta.$$

Hence the value of $J_2(\eta)$, defined in (4:8), for the set (5:1) is

$$(5:6) \quad Q(z) = z_\sigma z_\tau [\zeta_{\beta\sigma} \eta_{\beta\tau}(x_2) \eta_{\gamma\tau}(x_2) + \zeta_{\beta\tau} \eta_{\beta\sigma}(x_2)] \\ (\sigma, \tau = 1, \dots, t \leq n-1).$$

The inequality $t \leq n-1$ follows from the fact, noted in Theorem 4:2, that at least one of the $(2n+2)$ independent solutions of the canonical accessory equations has $\eta_1(x) = 0$. To determine the set (5:1) the $n+2$ conditions on the end values of the functions $\eta_1(x)$, given in (4:5), must be impressed upon the linear family, containing $2n+1$ or fewer parameters, including all the extremals which have the property that their functions $\eta_1(x)$ are not all identically zero. We now apply Theorem 4:2 and the results quoted in the preceding paragraphs to get the following theorem concerning the problems B and A.

THEOREM 5:1. Let E_{12} be a non-singular extremal arc for the problem A along which the second variation $J_2(\eta)$ in (4:8) is non-negative, and let the order of abnormality q of E be the same on every sub-interval of $x_1 x_2$. Then the two following conditions must hold:

(1) There can be no point 3 conjugate to the initial point 1 between 1 and 2 on the arc E_{12} , and the condition $Q(z) \geq 0$ must be satisfied for all values of the constants z not all zero, where

$$(5:7) \quad Q(z) = z_{\sigma} z_{\tau} [g_{\beta\gamma} \eta_{\beta\sigma}(x_2) \eta_{\gamma\tau}(x_2) - \lambda_{\beta\sigma} \eta_{\beta\tau}(x_2)] \\ (\beta, \gamma = 1, \dots, n; \sigma, \tau = 1, \dots, t \leq n-1).$$

In this expression

$$(5:8) \quad \eta_{\beta\sigma}(x) = \int_{x_1}^x [h_{\beta\gamma_0} \eta_{0\sigma} + h_{\beta\gamma_0'} \eta_{0\sigma}'] dx$$

and the functions $\eta_{0\sigma}(x)$ ($\sigma = 1, \dots, t$) form with the constants $\lambda_{\beta\sigma}$ a maximal set of solutions of the equations

$$(5:9) \quad d\omega_{\eta_0'} / dx - \omega_{\eta_0} + \lambda_{\beta} (dh_{\beta\gamma_0} / dx - h_{\beta\gamma_0'}) = 0, \\ \eta_{0\sigma}(x_1) = \eta_{0\sigma}(x_2) = 0,$$

for which the functions $\eta_{0\sigma}(x)$ are linearly independent. A con-
jugate point 3 to 1 on E_{12} is a point, other than 1, at which the
matrix

$$(5:10) \quad \left\| \begin{array}{c} \eta_{0\pi}(x_1) \\ \eta_{0\pi}(x_3) \\ \eta_{\beta\pi}(x_3) \end{array} \right\| \\ (\beta = 1, \dots, n; \pi = 1, \dots, t+2)$$

is of rank less than $t+2$, the set $\eta_{0\pi}$ ($\pi = 1, \dots, t+2$)
being a maximal set of solutions of the equation (5:9) for which
the functions $\eta_{0\pi}(x)$ are linearly independent, and the functions
 $\eta_{\beta\pi}$ being defined by equations analogous to (5:8).

(2) There can be no focal point 3 of the manifold $x_2 - a_2$
 $= y_0(x_2) = 0$ on E_{12} between its end points 1 and 2. The focal
points are defined by values x_3 for which the determinant

$$(5:11) \quad \left| \begin{array}{c} \eta_{0\rho}(x_3) \\ \eta_{0\rho}(x_2) \\ g_{\beta\gamma} [\eta_{\gamma\rho}(x_3) - \eta_{\gamma\rho}(x_2)] + \lambda_{\beta\rho} \end{array} \right| \\ (\beta, \gamma = 1, \dots, n; \rho = 1, \dots, n+2)$$

vanishes. Here the sets $\eta_{0\rho}(x)$, $\lambda_{\beta\rho}$ ($\rho = 1, \dots, n+2$) are a maximal set of linearly independent solutions of the equation (5:9), and the functions $\eta_{\gamma\rho}(x)$ are defined by equations analogous to (5:8).

According to Theorem 4:2 the columns of the determinant (2:20) form a complete set η_{1j} , ξ_{1j} ($i = 0, 1, \dots, n; j = 1, \dots, 2n+2$) of solutions of the canonical accessory equations (4:21). By deleting from that set the members linearly dependent upon sets for which $\eta_1(x) \equiv 0$, and by making use of the special nature of the set, one may readily relate the rank of the matrix (5:3) to that of the matrix (5:10) in such a way as to give the result stated in the theorem. To finish our discussion of the theorem we need to establish further only the fact that a focal point as defined in this section for problem B, or A, is a point at which the determinant (5:11) vanishes. To this end we desire to show that a value $x = x_3$ for which the determinant

$$(5:12) \quad \begin{vmatrix} \eta_{1j}(x_3) \\ \eta_{0j}(x_2) \\ \xi_{\beta\gamma} \eta_{\gamma j}(x_2) + \xi_{\beta j} \end{vmatrix}$$

($i = 0, \dots, n; j = 1, \dots, 2n+2; \beta, \gamma = 1, \dots, n$)

vanishes determines a focal point 3 of the manifold $x_2 - a_2 = y_0(x_2) = 0$. In this determinant the functions η_{1j} , ξ_{1j} are understood to form a complete set of solutions of the canonical accessory equations (4:21). First, let us suppose that the determinant (5:12) vanishes at a value $x_3 \neq x_2$. Then there exist constants a_j not all zero such that

$$\begin{aligned}
 (5:13) \quad & a_j \eta_{1j}(x_3) = 0, \\
 & a_j \eta_{0j}(x_2) = 0, \\
 & a_j [\xi_{\beta\gamma} \eta_{\gamma j}(x_2) + \zeta_{\beta j}] = 0.
 \end{aligned}$$

It follows that $\eta_1 = a_j \eta_{1j}$, $\zeta_1 = a_j \zeta_{1j}$ is a set of functions satisfying the conditions

$$(5:14) \quad \eta_0(x_2) = 0, \quad \xi_{\beta\gamma} \eta_{\gamma}(x_2) + \zeta_{\beta} = 0 \quad (\beta, \gamma = 1, \dots, n),$$

which, along with $\xi_2 = 0$, are the equations (5:4) and the accessory transversality condition. Further, η_1, ζ_1 so defined are such that the functions η_1 vanish at x_3 , but do not vanish identically on $x_3 x_2$. For if $\eta_1 \equiv 0$ on $x_3 x_2$ we have $\zeta_{\beta} = 0$ also, since the set $\eta_{\beta}, \zeta_{\beta}$ satisfies the second of equations (5:14).

One may verify by examining the canonical accessory equations

$$(4:21) \quad \text{that then } \zeta_0(x) \equiv 0 \text{ on } x_3 x_2. \text{ Hence we would have}$$

$\eta_1 = a_j \eta_{1j} \equiv 0$, $\zeta_1 = a_j \zeta_{1j} \equiv 0$ on $x_1 x_2$, which contradicts the assumption that the functions η_{1j}, ζ_{1j} form a complete set.

By definition the point 3 is then a focal point to the given manifold. Conversely, a value x_3 for which a non-identically zero set η_1 vanishes, where η_1, ζ_1 satisfy the equations (5:14), is a value for which the determinant (5:12) vanishes. This follows from the fact that such a set is expressible as a linear combination $\eta_1 = a_j \eta_{1j}$, $\zeta_1 = a_j \zeta_{1j}$ of the complete set η_{1j}, ζ_{1j} . For the constants a_j , not all zero, thus determined the equations (5:13) hold and therefore the determinant (5:12) vanishes. That the determinant (5:11) vanishes when and only when the determinant (5:12) vanishes may be seen by making use of the special properties of the complete set η_{1j}, ζ_{1j} yielded by the determinant (2:20). For, if we let $\eta_{1\rho}, \zeta_{1\rho}$ ($\rho = 1, \dots, n+2$) be a set linearly dependent on the set composed of the elements

of the first column and the last $n + 1$ columns of the determinant (2:20), we may write the determinant (5:12) in the form

$$\begin{vmatrix} 0 & \eta_{0\rho}(x_3) \\ & \eta_{\beta\rho}(x_3) \\ 0 & \eta_{0\rho}(x_2) \\ \xi_{\beta\alpha}\xi_{\alpha\gamma} & \xi_{\beta\gamma}\eta_{\gamma\rho}(x_2) + \zeta_{\beta\rho} \end{vmatrix}$$

$(\alpha, \beta, \gamma = 1, \dots, n; \rho = 1, \dots, n + 2;$
 $\xi_{\beta\beta} = 1; \xi_{\beta\gamma} = 0 \text{ if } \beta \neq \gamma).$

By substituting $\zeta_{\beta\rho} = -\lambda_{\beta\rho}$ and by suitably manipulating the rows of this determinant one derives the determinant (5:11).

As is readily seen from the preceding discussion, we have the following corollary to Theorem 5:1:

COROLLARY 5:1. The $n + 2$ sets of functions $\eta_{1\rho}$ and the associated constants $\lambda_{\beta\rho}$ appearing in the condition (2) of Theorem 5:1 may be taken to be the functions and constants

$$\begin{aligned} \eta_{1\gamma} &= y_{11\gamma}, & \eta_{1,n+1} &= y_{1a_0}, & \eta_{1,n+2} &= y_{1b_0}, \\ \lambda_{\beta\gamma} &= -\xi_{\beta\gamma}, & \lambda_{\beta,n+1} &= 0, & \lambda_{\beta,n+2} &= 0 \\ & (i = 0, 1, \dots, n; \beta, \gamma = 1, \dots, n), \end{aligned}$$

in which the right members are the derivatives occurring in the columns of the determinant (2:20). Suitable linear combinations of these derivatives may be used for the sets $\eta_{1\pi}$ and $\eta_{1\sigma}$, $\lambda_{\beta\sigma}$ occurring in condition (1).

6. Sufficiency theorems from the theory of the problem of Bolza. Sufficiency theorems for the problem A will be stated in terms of strengthened forms of the necessary conditions discussed in the preceding sections. We use the Roman numeral I to designate

the multiplier rule discussed in Section 2. Similarly, the symbol II_N' is used to designate the strengthened condition of Weierstrass, which is defined as follows: An arc E is said to satisfy the condition II_N' if the inequality

$$E(x, y_0, y_0', y_0'', \ell) > 0$$

holds at every element (x, y_0, y_0', ℓ) in a neighborhood N of those belonging to E and for all admissible sets $(x, y_0, y_0') \neq (x, y_0, y_0')$. For problem A the strengthened condition of Clebsch, denoted by III' , is the condition that the inequality

$h_{y_0' y_0''}(x, y_0, y_0', \ell) > 0$ shall hold along E . We define the condition IV' for problem A to be the condition (1) of Theorem 5:1 strengthened by the assumption that the quadratic form $Q(z)$ is not only positive but is positive definite. Likewise, the condition IV_F' is defined to be the condition (2) of Theorem 5:1 strengthened by excluding the possibility of a focal point 3 being at the point 1. We have here defined the strengthened conditions for problem A, or B. Where similar strengthened conditions are defined for problem C, we have the following results from the theory of the problem of Bolza [VII, p. 816, Theorem 9:4; X, p. 546, Theorems 1 and 3]:

Let E be an admissible arc without corners satisfying the equations $\phi_\beta = \psi_\mu = 0$ for a problem of Bolza, and satisfying further the conditions I, II_N' , III' , and either IV' or IV_F' with a set of multipliers $\ell_0 = 1, \ell_\beta(x), e_\mu$. Then there exists a neighborhood V of E in xy -space and a neighborhood T of the space of points $(x_1, y_{11}, x_2, y_{12})$ such that $J(C) > J(E)$ for every admissible arc C in V with ends in T satisfying the equations $\phi_\beta = \psi_\mu = 0$ and not identical with E .

When this result is applied to problem B and interpreted for problem A one obtains the following theorem:

THEOREM 6:1. SUFFICIENT CONDITIONS FOR A RESTRICTED STRONG RELATIVE MINIMUM. Let E be an admissible arc for the problem A such that E is without corners, joins two given fixed points in the xy-plane, and satisfies the conditions I, II_N', III', and either IV' or IV_P' with a set of multipliers $\ell_0 = 1, \ell_\beta$. Then there exists a neighborhood V of E such that $g(C) > g(E)$ for every admissible arc C in V not identical with E, joining the end points of E, and satisfying on x_1x_2 the condition that

$$(6:1) \quad \left| \int_{x_1}^{x_2} [h_\beta(C) - h_\beta(E)] dx \right| < \epsilon,$$

where ϵ is a sufficiently small positive constant.

In the theorem it is understood that the functions $h_\beta(C)$, $h_\beta(E)$ are the functions $h_\beta(x, y, y')$ occurring in the equations (1:1), with arguments the functions y, y' defining C and E respectively. Similar notations are used in writing $I_\beta(E)$, $I_\beta(C)$, $g(E)$ and $g(C)$.

It is desirable to weaken the restrictive condition (6:1). In the next section we show that with an additional normality assumption we may weaken the condition (6:1) to the extent that it be required to hold only at $x = x_2$. Reflection on the possible nature of the function g shows that one could not in general remove the condition entirely.

The following sufficiency theorem for the accessory minimum problem, which was discussed in Section 4, will be found useful in the next section:

THEOREM 6:2. Let E have the properties specified in Theorem 6:1 except possibly the property that it satisfies II_N'. Then along E the inequality $J_2(\eta) > 0$ holds for every admissible

variation $\eta_0(x)$ satisfying the equations $\eta_0(x_1) = \eta_0(x_2) = 0$ and not identically zero on x_1x_2 .

For the proof of the theorem we shall refer to the literature of the problem C of Bolza. Results for problem A are then immediate by the use of Theorem 4:1. For the case when E satisfies IV' we refer to two theorems by Hestenes [VII, p. 811, Theorem 8:3, and X, p. 546, Theorem 3]. When E satisfies IV_F' we refer to three lemmas, again by Hestenes [X, pp. 549-550, Lemmas 8, 5, 6].

We note that the accessory minimum problem for problem A is a problem of the type A, since the functional to be minimized is a function of integrals.

7. Extension of the sufficiency proof. We have seen that the theory for the problem of Bolza yields a restricted relative minimum for problem A, and we noted that a weakening of the restrictive condition (6:1) on the admissible arcs is desirable. We may effect the desired weakening of that condition if E satisfies an additional normality condition, namely, that the arc E satisfies the Euler-Lagrange equation $dh_{y_0}/dx - h_{y_0} = 0$ with only one set of multipliers when one keeps a non-zero multiplier fixed. Subarcs of E need not satisfy the normality condition. The theorem to be proved in this section may now be formulated as follows:

THEOREM 7:1. SUFFICIENT CONDITIONS FOR A STRONG RELATIVE MINIMUM. If the admissible arc E satisfies the hypotheses of Theorem 6:1 and in addition the normality condition defined above, then the conclusions of Theorem 6:1 hold with the restrictive condition (6:1) required to hold only at $x = x_2$.

The proof of the theorem will be based on a series of five lemmas, the last three of which are similar to lemmas used by Hestenes to prove an analogous result for the isoperimetric problem [XI]. In these lemmas it is understood that the arc E satisfies the hypotheses made in Theorem 7:1. As a preliminary to the first lemma we note that in Section 1 it was assumed that not all the first partial derivatives of g vanish on E . Without loss of generality we may assume that $g_1 \neq 0$. In fact we may suppose that $g_1 > 0$ since we may replace I_1 by $-I_1$ if necessary. We shall fix the multiplier $\lambda_1 = g_1(E)$. With these statements in mind we formulate the following lemma:

LEMMA 7:1. With the fixed relation $\lambda_1 = g_1(E)$ the normality condition is equivalent to the condition that the equation

$$(7:1) \quad [dh_{py_0}/dx - h_{py_0}]a_p = 0 \quad (p = 2, \dots, n)$$

holds along E only in case the constants a_p are all zero.

For, if two distinct sets of multipliers $\lambda_1 = g_1(E)$, λ_p ($p = 2, \dots, n$) and $\lambda_1 = g_1(E)$, λ_p satisfy the Euler-Lagrange equation with E so also does the non-null set $\lambda_1 - \lambda_1 = 0$, $\lambda_p - \lambda_p$. Then equation (7:1) holds with $a_p = \lambda_p - \lambda_p$. Conversely, if a_p is a non-null set for which (7:1) holds on E , then the Euler-Lagrange equation is satisfied on E by the two distinct sets $\lambda_1 = g_1(E)$, λ_p and $\lambda_1 = g_1(E)$, $\lambda_p + a_p$, where λ_1 , λ_p are the multipliers belonging to E .

LEMMA 7:2. If $\eta_0(x') = \eta_0(x'') = 0$ then

$$(7:2) \quad \lambda_p [\eta_p(x'') - \eta_p(x')] = 0 \quad (x_1 \leq x' \leq x'' \leq x_2)$$

holds for every admissible variation $\eta_1(x)$. From this it follows that if $\lambda_1 \neq 0$, $\eta_1(x_1) = \eta_p(x_1) = \eta_p(x_2) = 0$ ($p = 2, \dots, n$), then $\eta_1(x_2) = 0$ also.

That the equation (7:2) holds under the conditions stated in the lemma is readily provable by multiplying the equations of variation

$$h_{\beta y_0} \eta_0 + h_{\beta y_0'} \eta_0' - \eta_{\beta}' = 0 \quad (\beta = 1, \dots, n)$$

by the multipliers ℓ_{β} , adding, and applying the usual integration by parts with the help of the equation $dh_{y_0'}/dx - h_{y_0} = 0$.

In the next lemmas there appears an integral J_{ℓ} defined by the equation

$$(7:3) \quad J_{\ell} = \int_{x_1}^{x_3} h(x, y_0, y_0', \ell) dx,$$

in which the integrand is the sum $\sum \ell_{\beta} h_{\beta}$. A value $x_3 \neq x_1$ is said to define a point 3 conjugate to 1 on E relative to J_{ℓ} if there exists a solution $\eta(x)$ of the equation

$$d\omega_{\eta}/dx - \omega_{\eta} = 0,$$

having $\eta(x_1) = \eta(x_3) = 0$ and $\eta \neq 0$ on $x_1 x_3$, where the multipliers ℓ_{β} appearing in J_{ℓ} and ω are those belonging to E. The function ω that occurs here is defined by equation (4:7).

LEMMA 7:3. HAHN'S LEMMA. If E has on it no point 3 conjugate to 1 relative to J_{ℓ} , there exist neighborhoods V_1 of E in xy-space, N of the end values of E in $(x_1 y_1 x_2 y_2)$ -space and L of the multipliers ℓ_{β} belonging to E such that for every set $(x_1 y_1 x_2 y_2)$ in N and ℓ_{β} in L there is an extremal E_{ℓ} in V_1 having $(x_1 y_1 x_2 y_2)$ as its end values and ℓ_{β} as its multipliers. Moreover, for these values of ℓ_{β} the inequality $J_{\ell}(C) > J_{\ell}(E_{\ell})$ holds for every admissible arc C in V_1 joining the ends of E_{ℓ} and not identical with E_{ℓ} .

This lemma is analogous to one given by Hahn and has been established by Birkhoff and Hestenes [IX, pp. 253-254] following a method given by Bliss [IV, pp. 267-270] in the proof of a

similar theorem for the problem of Bolza. The assumptions made in Theorem 7:1 do not exclude the possibility of there being pairs of conjugate points on E relative to J_ℓ . However, the arc E can be subdivided into a finite number of subarcs on each of which there are no pairs of conjugate points relative to J_ℓ [cf. III, p. 739]. In this case we have the following extension of Lemma 7:3:

LEMMA 7:4. EXTENDED LEMMA OF HAHN. Let $t_0 = x_1 < t_1 < \dots < t_{w+1} = x_2$ be a set of values such that there is no pair of conjugate points on E relative to J_ℓ on any of the intervals $t_{r-1} \leq x \leq t_r$ ($r = 1, \dots, w + 1$). Then there exist neighborhoods V_1 of E in xy -space, M of the values $(x, y) = (t_s, b_s)$ ($s = 1, \dots, w$) belonging to E and L of the multipliers ℓ_β belonging to E such that for every set (t_s, b_s) in M and ℓ_β in L there is a broken extremal E_ℓ in V_1 joining the points 1 and 2, passing through the points (t_s, b_s) , having no corners on the intervals $t_{r-1} < x < t_r$, and having ℓ_β as its multipliers. Moreover, for these values of ℓ_β the inequality $J_\ell(C) > J_\ell(E_\ell)$ holds for every admissible arc C in V_1 joining the end points of E_ℓ , passing through the points (t_s, b_s) and not identical with E_ℓ .

Lemma 7:4 follows as a result of applying Lemma 7:3 to the successive subarcs of E determined by the intervals $t_{r-1} \leq x \leq t_r$ ($r = 1, \dots, w + 1$).

LEMMA 7:5. Let $t_0, t_1, \dots, t_{w+1}$ be the set of values defined in Lemma 7:4. There exists a $(w + n - 1)$ -parameter family of broken extremals

$$\begin{aligned}
 & y(x, b_1, \dots, b_w, c_2, \dots, c_n), \\
 (7:4) \quad & \ell_1 = g_1(E), \quad \ell_p(b_1, \dots, b_w, c_2, \dots, c_n) \\
 & (p = 2, \dots, n; x_1 \leq x \leq x_2)
 \end{aligned}$$

containing E for $b_s = b_{s0}$, $c_p = c_{p0}$, passing through the points 1, 2 and $(x, y) = (t_s, b_s)$, having no corners on the intervals $t_{r-1} < x < t_r$, and satisfying the conditions

$$(7:5) \quad \int_{x_1}^{x_2} h_p(x, y, y') dx = c_p.$$

The functions $y(x, b, c)$, $y_x(x, b, c)$, $l_p(b, c)$ have continuous first and second derivatives for values (x, b, c) near those belonging to E except possibly at the corner points.

In order to establish this lemma, let

$$(7:6) \quad Y(x, b_1, \dots, b_w, l_2, \dots, l_n), \quad l_1 = g_1(\bar{E}), \quad l_p \quad (x_1 \leq x \leq x_2)$$

be a $(w + n - 1)$ -parameter family of broken extremals containing E for values $b_s = b_{s0}$, $l_p = l_{p0}$, passing through the points 1, 2 and $(x, y) = (t_s, b_s)$, and having no corners on the intervals $t_{r-1} < x < t_r$. Except at corner points the functions $Y(x, b, l)$ $Y_x(x, b, l)$ have continuous first and second derivatives for values (x, b, l) near those belonging to E. The existence of a family of this type follows readily from the first part of Lemma 7:4 and our earlier studies, in Section 2, of the properties of functions defining extremals. When the function $Y(x, b, l)$ is substituted in the integrals (7:5) a set of functions $I_p(b, l)$ is obtained having continuous first and second derivatives for values (b, l) near those on E. The functional determinant $|\partial I_p / \partial l_q|$ ($p, q = 2, \dots, n$) is different from zero when $(b, l) = (b_0, l_0)$, as will be seen in the next paragraph. The equations $I_p(b, l) = c_p$ are satisfied by the values $(b, l) = (b_0, l_0)$, and hence have unique solutions $l_p(b, c)$ with $l_p(b_0, c_0) = l_{p0}$ and having continuous second derivatives near $(b, c) = (b_0, c_0)$. When the functions $l_p(b, c)$ are substituted for l_p in the functions (7:6) a family (7:4) is obtained having the properties described

in the lemma.

To complete the proof of Lemma 7:5 we have yet to show that the determinant $|\partial I_p / \partial \ell_q|$ is different from zero at $(b, \ell) = (b_0, \ell_0)$. First, we notice that no variation of the form

$\eta = Y_{\ell_p}(x, b_0, \ell_0) a_p$ with the set $(a_p) \neq (0)$ is identically zero. For, if we substitute the functions (7:6) in the equation $dh_{y_0}/dx - h_{y_0} = 0$, differentiate with respect to ℓ_p , multiply by a_p , and sum, we find that along E the equation

$$(7:7) \quad d\omega_{\eta}/dx - \omega_{\eta} + [dh_{py_0}/dx - h_{py_0}]a_p = 0$$

holds, where ω is the function (4:7) and its arguments are

$\eta = Y_{\ell_p} a_p$, $\eta' = Y_{\ell_p}' a_p$. If $\eta \equiv 0$, then equations (7:1) would hold with a set $(a) \neq (0)$, which is impossible according to Lemma 7:1. Next we notice that if $|\partial I_p / \partial \ell_q| = 0$, then there exist constants a_q not all zero such that the equations

$$(7:8) \quad a_q \partial I_p / \partial \ell_q = \int_{x_1}^{x_2} (h_{py_0} \eta + h_{py_0}' \eta') dx = 0$$

hold, where $\eta = Y_{\ell_q} a_q$. Multiplying the left member of (7:7) by this η , integrating from x_1 to x_2 and using (7:8), we find that

$$(7:9) \quad \int_{x_1}^{x_2} (\eta d\omega_{\eta}/dx - \omega_{\eta} \eta) dx = 0.$$

With the usual integration by parts this becomes

$$(7:10) \quad \sum_{r=1}^{w+1} \omega_{\eta} \eta \Big|_{t_{r-1}}^{t_r} + \int_{x_1}^{x_2} 2\omega dx = 0.$$

By differentiating with respect to ℓ_q the identities $Y(t_v, b, \ell) = b_v$ ($v = 0, 1, \dots, w+1$), in which $b_0 = y_{01}$, $b_{w+1} = y_{02}$, we find that $Y_{\ell_q}(t_v, b, \ell) = 0$. Therefore $\eta(t_v) = Y_{\ell_q}(t_v, b_0, \ell_0) a_q = 0$, and (7:10) reduces to

$$(7:11) \quad \int_{x_1}^{x_2} 2\omega dx = 0.$$

In view of equations (7:8) the functions

$$\eta_\beta(x) = \int_{x_1}^x (h_{\beta y_0} \eta + h_{\beta y_0'} \eta') dx \quad (\beta = 1, \dots, n)$$

satisfy the relations $\eta_1(x_1) = \eta_p(x_1) = \eta_p(x_2) = 0$ ($p = 2, \dots, n$). Hence $\eta_1(x_2) = 0$ also, by Lemma 7:2. It follows that the terms $g_{\beta\gamma} \eta_\beta(x_2) \eta_\gamma(x_2)$ outside the integral sign in the expression (4:8) for $J_2(\eta)$ vanish so that $J_2(\eta)$ reduces to the left member of (7:11). But then the equality (7:11) contradicts Theorem 6:2. Therefore our assumption that $|\partial I_p / \partial \eta_q| = 0$ is contrary to fact, and Lemma 7:5 is established.

We may now establish Theorem 7:1. First, let V and ϵ be chosen so that Theorem 6:1 holds, and let V_1 , M , and L be chosen so that Lemma 7:4 holds. We diminish M and L , if necessary, so that the broken extremals E_λ of Lemma 7:4 satisfy the condition (6:1). That we may do this follows from the continuity properties of the family of broken extremals given in (7:4). We now let V_2 be a neighborhood of E in xy -space interior to V and V_1 and such that points $(x, y) = (t_s, b_s)$ of Lemma 7:4 lying in V_2 are also in M . Diminish ϵ if necessary so that the inequality $g_1(C) > 0$ holds for each curve C satisfying at $x = x_2$ the condition (6:1). This we may do since we have assumed that $g_1(E) > 0$. The region $V = V_2$ and the constant ϵ so chosen will be shown to have the property described in Theorem 7:1. To establish this we consider a given admissible arc C lying in V_2 joining the end points of E , and satisfying at $x = x_2$ the condition (6:1). By choosing the parameters b_s equal to the ordinates of C at t_s and the parameters c_p equal to the values of the integrals I_p evaluated along C we determine, as in Lemma 7:5, a broken extremal E_λ of the family

(7:4) passing through the points (t_s, b_s) and having $I_p(E_\ell) = I_p(C)$ ($p = 2, \dots, n$). By Lemma 7:4 we have

$$0 < J_\ell(C) - J_\ell(E_\ell) = \ell_\beta [I_\beta(C) - I_\beta(E_\ell)] = \ell_1 [I_1(C) - I_1(E_\ell)]$$

if $C \neq E$. Since $\ell_1 = g_1(E) > 0$ it follows that $I_1(C) > I_1(E_\ell)$ unless $C \equiv E_\ell$. By Taylor's formula we have the relation

$$g(C) - g(E_\ell) = [I_1(C) - I_1(E_\ell)]g_1,$$

where the arguments of g_1 are $I_1(E_\ell) + \theta[I_1(C) - I_1(E_\ell)]$ ($0 < \theta < 1$), and $I_p(E_\ell)$. But $g_1 > 0$ with these arguments by virtue of our choice of ϵ . Hence $g(C) > g(E_\ell)$ ($C \neq E_\ell$). By our choice of V_2 and ϵ the broken extremal E_ℓ satisfies the condition (6:1). We may then apply Theorem 6:1 to derive the result $g(E_\ell) \geq g(E)$, the equality holding only if $E_\ell \equiv E$. It follows that $g(C) > g(E)$ ($C \neq E$). Theorem 7:1 has now been established.

8. Application to a special problem. In the introduction we noted that Bolza had treated in detail two problems discussed by Euler. Bolza generalized the problems formulated by Euler and by means of special methods gave sufficiency proofs for a strong relative minimum for one and maximum for the other. Both of the problems are problems of the type A formulated in Section 1. In this section we shall apply the theory developed for problem A to one of the problems treated by Bolza.

The first of the two problems considered by Bolza may be formulated as an instance of problem A as follows: In a class of admissible arcs joining the two fixed points $(x_1, y_1) = (-a, 0)$, $(x_2, y_2) = (a, 0)$ ($a > 0$) we seek an arc minimizing the function

$$(8:1) \quad g(I_1, I_2) = I_2 / (I_1 + b^2),$$

in which b^2 is a positive constant and I_1 , I_2 are the area and length integrals

$$(8:2) \quad I_1 = \int_{-a}^a y \, dx, \quad I_2 = \int_{-a}^a \sqrt{1 + y'^2} \, dx.$$

This problem may be interpreted geometrically in the following manner: Let b^2 represent the area of the rectangle with corners at the points $(-a, 0)$, $(-a, -b^2/2a)$, $(a, -b^2/2a)$, $(a, 0)$. Then in a class of admissible arcs joining the two upper corners of the rectangle and satisfying the relation $y(x) > -b^2/2a$ ($-a < x < a$) we seek an arc minimizing the ratio of the length of the arc to the area bounded by the arc and the base and sides of the above specified rectangle. We then take the region R of Section 1 to be the set of all points (x, y, y') having $y > -b^2/2a$.

On applying Theorem 2:1, that is, the multiplier rule, to the above special problem we find that a minimizing arc E for the problem must satisfy the equation

$$(8:3) \quad \ell_2 y' / \sqrt{1 + y'^2} = \ell_1 (x - x_1) + c,$$

in which ℓ_1 , ℓ_2 are constant multipliers which satisfy with E the transversality condition

$$(8:4) \quad \ell_1 [I_1 + b^2]^2 = -I_2, \quad \ell_2 [I_1 + b^2] = 1.$$

We have set $\ell_0 = 1$ in these equations. Solutions of (8:3) are functions defining arcs of circles and satisfying the equation

$$(8:5) \quad (x + h)^2 + (y + k)^2 = (\ell_2 / \ell_1)^2,$$

where h , k are constants. The branches of the circles (8:5) which pass through $(-a, 0)$, $(a, 0)$ and which are single valued on $-a \leq x \leq a$ are given by the formula

$$(8:6) \quad y = \pm(-k + \sqrt{R^2 - x^2}),$$

where $k > 0$ and $R = -l_2/l_1$. A minus sign appears in the last relation because $l_2 > 0$, $l_1 < 0$ for those of the arcs (8:6) which also satisfy equations (8:4). We calculate I_1 , I_2 along the arcs (8:6) and then make the substitution $\omega = 2 \sin^{-1}(a/R)$. The result is

$$(8:7) \quad \begin{aligned} I_1 &= -a^2 \cot \omega/2 + a^2(\omega/2) \operatorname{cosec}^2 \omega/2, \\ I_2 &= a \omega \operatorname{cosec} \omega/2. \end{aligned}$$

By dividing the members of one of the equations (8:4) by the respective members of the other and then substituting the values (8:7) for I_1 , I_2 and $a \operatorname{cosec} \omega/2$ for $(-l_2/l_1)$ one derives the equation (8:8) appearing in the following lemma:

LEMMA 8:1. Circular arcs defined by the equation (8:6), joining the two fixed points $(-a, 0)$, $(a, 0)$ and subtended by a central angle ω satisfy the transversality condition if and only if ω is a solution of the equation

$$(8:8) \quad b^2(1 - \cos \omega) - a^2(\omega + \sin \omega) = 0.$$

The equation (8:8) has real roots when and only when $b^2/a^2 \geq \pi/2$. If this relation holds without the equality sign one of the roots lies between 0 and π , the other between π and 2π . In this case let E_0 designate the arc of the family (8:6) whose central angle ω is the smaller of the two roots of (8:8). Then E_0 satisfies the condition I for our special problem.

Bolza derived equation (8:8) by another method [II, pp. 249-251]. He established the readily-verified statement occurring in the lemma concerning the roots of the equation (8:8). We shall show in the succeeding paragraphs that besides satisfying the condition I, E_0 satisfies also the conditions II_N' , III' ,

IV' and the normality condition, so that we may apply Theorem 7:1 to establish a strong relative minimum for the function (8:1) along E_0 . It will appear later that the arc determined by the larger of the roots of the equation (8:8) does not satisfy the relation $Q(z) > 0$. A similar statement applies when the two roots are equal, i.e., when $b^2/a^2 = \pi/2$. For that reason we limit our discussion to E_0 .

The conditions II_N' and III' are strengthened forms of the conditions stated in Theorem 3:1. The function h appearing in Theorem 3:1 is for our special problem

$$(8:9) \quad h = \ell_1 y + \ell_2 \sqrt{1 + y'^2}.$$

Setting this value of h in the function $E(x, y_0, y_0', y_0', \ell)$ defined in (3:4) we find

$$(8:10) \quad \begin{aligned} E &= \ell_2 [\sqrt{1 + Y'^2} - \sqrt{1 + y'^2} - (Y' - y')y' / \sqrt{1 + y'^2}] \\ &= \ell_2 [\sqrt{(Y'y' + 1)^2 + (Y' - y')^2} - \sqrt{(Y'y' + 1)^2}] / \sqrt{1 + y'^2}. \end{aligned}$$

Examination of the second equation in (8:4) shows that $\ell_2 > 0$. We readily see then that the condition II_N' is satisfied by E_0 . Since the function $h_{y', y'}(x, y, y', \ell)$ occurring in the condition III' is here

$$(8:11) \quad h_{y', y'} = \ell_2 / (1 + y'^2)^{3/2},$$

we see that along E_0 the condition III' is also satisfied. We summarize the results of this paragraph in the following lemma:

LEMMA 8:2. The conditions II_N' and III' are satisfied along E_0 . The Weierstrass E function for the special problem is given in (8:10), and the function appearing in the Clebsch condition is given in (8:11).

For our special problem the matrix (5:10) defining pairs of conjugate points contains 4 rows and at most 3 columns since for this problem $n = 2$ and the number of columns is $t + 2 \leq n + 1$ by (3:7). Corollary 5:1 states that the functions appearing in the matrix (5:10) may be taken from the determinant (2:10), in which the elements are derivatives of the functions y_1, z_1 taken with respect to parameters which are then evaluated on E_0 . Accordingly, we differentiate the function $y = -k + \sqrt{k^2 - (x + h)^2}$, which is a part of the family (3:5) containing E_0 , with respect to the constants k, h, k and set $h = 0$ to obtain the 3 functions

$$(5:11) \quad \eta_{01} = 1, \quad \eta_{02} = -x/\sqrt{k^2 - x^2}, \quad \eta_{03} = k/\sqrt{k^2 - x^2}.$$

These 3 functions satisfy the relation $\eta_{0\pi}(x)a_\pi \leq 0$ ($\pi = 1, 2, 3$) for all sets $(a) \in (0)$. We have seen that there is a set of at most 3 such functions, so this set is a maximal set. The corresponding constants $\lambda_{\beta\pi}$ ($\beta = 1, 2; \pi = 1, 2, 3$), needed for the calculation of $Q(2)$, are found by differentiating $z_1 = -l_1$, $z_2 = -l_2 k$ with respect to k, h, k and then setting the results equal to $-L_{a\pi} = -\lambda_{\beta\pi}$. In that manner we find

$$(5:12) \quad \lambda_{1\pi} = \lambda_{21} = \lambda_{22} = 0, \quad \lambda_{23} = -l_1.$$

To simplify further computations we make the substitution $\theta = \sin^{-1} x/k$. From (5:12) we have then

$$(5:13) \quad \eta_{01} = 1, \quad \eta_{02} = -\tan \theta, \quad \eta_{03} = \sec \theta.$$

By use of equations similar to (5:8), in which $h_1 = y$, $h_2 = \sqrt{k^2 - x^2}$, the functions $\eta_{\beta\pi}$ appearing in (5:10) are found to be

$$(5:14) \quad \begin{aligned} \eta_{11} &= -k(\sin \theta - \sin \theta_1), & \eta_{21} &= 0, \\ \eta_{12} &= k(\cos \theta - \cos \theta_1), & \eta_{22} &= \sec \theta - \sec \theta_1, \\ \eta_{13} &= k(\theta - \theta_1), & \eta_{23} &= [\tan \theta_1 - \tan \theta + \theta - \theta_1], \end{aligned}$$

where $\theta_1 = -\sin^{-1}a/R$. The matrix (5:10) determining conjugate points is then

$$(8:16) \begin{vmatrix} 1 & \tan \theta & \sec \theta \\ 1 & \tan \theta_1 & \sec \theta_1 \\ R(\sin \theta - \sin \theta_1) & R(\cos \theta_1 - \cos \theta) & R(\theta - \theta_1) \\ 0 & \sec \theta_1 - \sec \theta & \tan \theta_1 - \tan \theta + \theta - \theta_1 \end{vmatrix}$$

where the signs have been changed in the first two columns. In order to show that this matrix has rank 3 on $\theta_1 < \theta < \theta_2$, corresponding to $x_1 < x < x_2$, one may demonstrate that the determinant composed of the first, second and fourth rows has the value $2 \sin(\theta - \theta_1)[\tan(\theta - \theta_1)/2 - (\theta - \theta_1)/2]/\cos \theta \cos \theta_1$. This expression does not vanish on the interval $\theta_1 < \theta < \theta_2 = -\theta_1$, provided $\theta_2 < \pi/2$. We note that $\omega/2 = \theta_2$ by definition, and that E_0 was determined by $\omega < \pi$. Therefore $\theta_2 < \pi/2$, and it follows that the matrix (8:16) has rank 3 on the interval $\theta_1 < \theta < \theta_2$. There is then no point 3 conjugate to 1 on $-a < x < a$, which is the interval corresponding to $\theta_1 < \theta < \theta_2$. We may note here that the range of θ is limited to the interval $-\pi/2, \pi/2$ by the requirement that admissible arcs be defined by single-valued functions $y(x)$. We have now established part of the following lemma:

LEMMA 8:3. The arc E_0 satisfies the condition IV'. A matrix determining points conjugate to 1 is the matrix (8:16) which has rank 3 on $\theta_1 < \theta < \theta_2$ or $x_1 < x < x_2$. The function $Q(z)$ defined by (5:7) has the value

$$(8:17) \quad Q(z) = 2 \ell_1(\theta_2 - \tan \theta_2)z^2$$

on E_0 and satisfies the inequality $Q(z) > 0$ if $z \neq 0$.

To complete the proof of this lemma we still must show that $Q(z) > 0$ holds on E_0 . For the calculation of $Q(z)$ we need a linear combination of the set (8:14) which satisfies $\eta_0(\theta_1) = \eta_0(\theta_2) = 0$. Such a combination is $\eta_0 = \sec \theta - \sec \theta_2$. Corresponding linear combinations of the sets in (8:15) and (8:13) with $\theta_1 = -\theta_2$ yield

$$(8:18) \quad \begin{aligned} \eta_1(x_2) &= 2R(\theta_2 - \tan \theta_2), & \lambda_1 &= 0, \\ \eta_2(x_2) &= 2(\theta_2 - \tan \theta_2), & \lambda_2 &= -\lambda_1. \end{aligned}$$

For this set $Q(z)$ has the value

$$(8:19) \quad Q(z) = z^2 [4(g_{11}R^2 + 2g_{12}R)(\theta_2 - \tan \theta_2)^2 + 2\lambda_1(\theta_2 - \tan \theta_2)].$$

By use of the relations (8:7), (8:8) and a cosec $\theta_2 = R$, $\omega = 2\theta_2$, one may show that along E_0

$$(8:20) \quad I_1 + b^2 = 2\theta_2 R^2, \quad I_2 = 2\theta_2 R.$$

These equations may then be combined with the equation (8:19) and the definition of g given in (8:1) to yield the value of $Q(z)$ appearing in (8:17). Since $\lambda_1 < 0$, $Q(z) > 0$ holds provided $0 < \theta_2 < \pi/2$. This last relation implies $b^2/a^2 > \pi/2$, and conversely, as we noted earlier. Lemma 8:3 has now been established.

One readily verifies that the equations

$$dh_{1y}/dx - h_{1y} = 0, \quad dh_{2y}/dx - h_{2y} = 0$$

do not hold along E_0 , and that therefore E_0 satisfies the normality condition defined in Section 7. By virtue of this statement and Lemmas 8:1, 8:2 and 8:3, the arc E_0 satisfies all the hypotheses of Theorem 7:1, and the following theorem is proved:

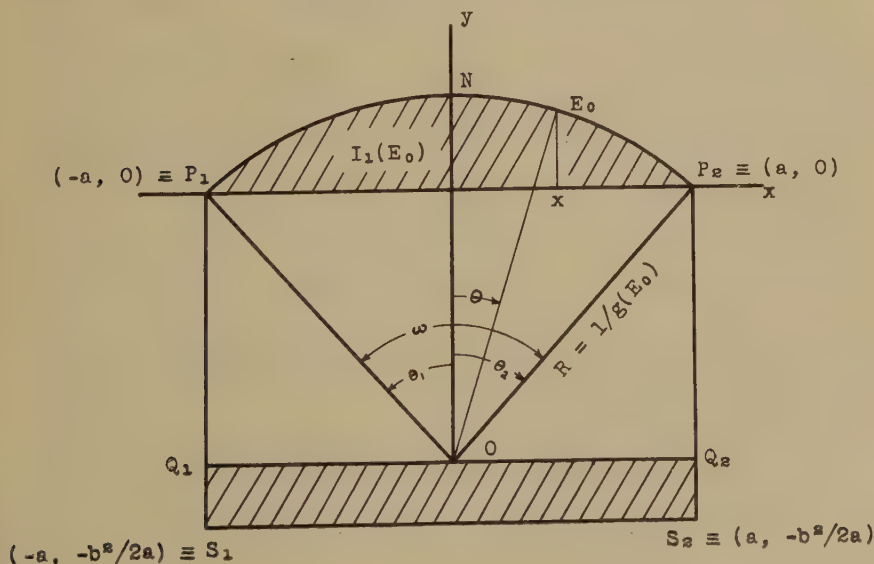
THEOREM 8:1. For the special problem considered in this section there exists an extremal E_0 defined as in Lemma 8:1 such

that the function g defined in (8:1) takes on a strong relative minimum on E_0 provided only that $b^2/a^2 > \pi/2$.

The transversality condition for the special problem may be interpreted geometrically. Euler [I, p. 154] showed that the minimum value of g is equal to the curvature of the minimizing arc. We may verify this by dividing the members of the second of equations (8:20) by the respective members of the first of the equations. That division yields $g(E_0) = 1/R$, where R is the radius of the circle of which E_0 is an arc. The first of the equations (8:20) further shows that on E_0

$$(8:21) \quad b^2 = 2\theta_2 R^2 - I_1 = \omega R^2 - I_1,$$

or the constant b^2 is equal to twice the area of the sector subtended by ω less the area I_1 between the arc E_0 and the chord $(-a, 0), (a, 0)$. Euler also expressed this fact, though in a somewhat different manner. The above relations and the relations between ω , θ , and x are shown in the figure which follows.



If we represent an area by a succession of letters designating its perimeter then by (8:21) we have

$$\begin{aligned} b^2 &= 2OP_1NP_2 - I_1(E_0) = 2OP_1P_2 + I_1(E_0) \\ &= P_1Q_1Q_2P_2 + I_1(E_0). \end{aligned}$$

But $b^2 = P_1S_1S_2P_2$ so that $I_1(E_0) = Q_1S_1S_2Q_2$. Thus we see that the center O of the minimizing arc E_0 must lie far enough above S_1S_2 to make the area of the rectangle $Q_1S_1S_2Q_2$ lying immediately below O equal to the area $I_1(E_0)$. A consequence of this statement is that the y coordinate of O must satisfy the relation $y > -b^2/2a$.

9. A second application. By replacing the function g in (8:1) by the function

$$(9:1) \quad g(I_1, I_2) = I_2[I_1 + b^2]$$

we have a statement in the first part of Section 8 of the second problem considered by Bolza. The discussion of this second problem may be made to follow the discussion given in Section 8 for the first problem with a few changes of formulas and results. We shall note briefly most of the necessary changes. The transversality condition (8:4) should be replaced by $l_1 = I_2$, $l_2 = I_1 + b^2$. This condition leads to the equation

$$(9:2) \quad b^2(1 - \cos \omega) + a^2(3\omega - \sin \omega) = 0$$

rather than (8:8). This equation has distinct real roots only when $b^2/a^2 > 1/x$, where

$$(9:3) \quad x = \frac{\sin 2\gamma}{3 - \cos 2\gamma} = .2785 \text{ approx.,}$$

the angle γ occurring here being the root of

$$(9:4) \quad 3\gamma = 2 \tan \gamma$$

lying between 0 and $\pi/2$. Approximately $\gamma = 55^\circ 26'$. The equations (9:2), (9:3) and (9:4) occur in Bolza's treatment of this problem [II, pp. 254-255]. In Lemma 8:1 we determine E_0 for the second problem by the root of (9:2) lying between 0 and -2γ . Then $\theta_2 = \omega/2$ lies between 0 and $-\gamma$. With this value of θ_2 determining the range of θ , $\theta_1 \geq \theta \geq \theta_2 = -\theta_1$, one may readily verify as before that the rank of the matrix (8:16) is 3 on $\theta_2\theta_1$, or that E_0 satisfies the conjugate point condition. Instead of (8:19) one has

$$Q(z) = z^2 [8R(\theta_2 - \tan \theta_2)^2 + 2\ell_1(\theta_2 - \tan \theta_2)].$$

By use of the equations

$$(9:5) \quad I_1 + b^2 = -2\theta_2 R^2 = \ell_2, \quad I_2 = 2\theta_2 R = \ell_1,$$

which correspond to the equations (8:20), one may reduce $Q(z)$ to

$$(9:6) \quad Q(z) = 4Rz^2(\theta_2 - \tan \theta_2)(3\theta_2 - 2 \tan \theta_2),$$

corresponding to (8:17). The values $R = a \operatorname{cosec} \omega/2$ and θ_2 occurring here are negative. The relation $Q(z) > 0$ then holds for $-\gamma < \theta_2 < 0$, where γ has the value defined above. We may note that $Q(z) = 0$ when $\theta_2 = -\gamma$. The Weierstrass and Clebsch conditions of Lemma 8:2 are satisfied along E_0 since $\ell_2 > 0$, as is shown by the equations (9:5). Corresponding to Theorem 8:1 we have the following theorem:

THEOREM 9:1. For the second problem considered by Bolza and discussed above there exists an extremal E_0 such that the function g defined in (9:1) takes on a strong relative minimum on E_0 , provided only that $b^2/a^2 > 1/\chi$, where χ has the value given in (9:3).

BIBLIOGRAPHY

- I. Euler, L., Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, Lausanne et Genève (1744).
- II. Bolza, O., Two propositions of Euler's from the calculus of variations, Annali di Matematica, vol. 20 (1913), pp. 245-255.
- III. Bliss, G. A., The problem of Lagrange in the calculus of variations, American Journal of Mathematics, vol. 52 (1930), pp. 673-744.
- IV. Bliss, G. A., The problem of Bolza in the calculus of variations, Annals of Mathematics, ser. 2, vol. 33 (1932), pp. 261-274.
- V. Bliss, G. A., and Hestenes, M. R., Sufficient conditions for a problem of Mayer in the calculus of variations, Transactions of the American Mathematical Society, vol. 35 (1933), pp. 305-326.
- VI. Hestenes, M. R., Sufficient conditions for the general problem of Mayer with variable end-points, Transactions of the American Mathematical Society, vol. 35 (1933), pp. 479-490.
- VII. Hestenes, M. R., Sufficient conditions for the problem of Bolza in the calculus of variations, Transactions of the American Mathematical Society, vol. 36 (1934), pp. 793-818.
- VIII. Bliss, G. A., The problem of Bolza in the calculus of variations, Mimeographed lecture notes, The University of Chicago, 1935.
- IX. Birkhoff, G. D., and Hestenes, M. R., Natural isoperimetric conditions in the calculus of variations, Duke Mathematical Journal, vol. 1 (1935), pp. 251-258.
- X. Hestenes, M. R., On sufficient conditions in the problems of Lagrange and Bolza, Annals of Mathematics, ser. 2, vol. 37 (1936), pp. 543-551.
- XI. Hestenes, M. R., A sufficiency proof for isoperimetric problems in the calculus of variations, To appear in the 1938 issue of the Bulletin of the American Mathematical Society.

VITA

Carroll Parker Brady was born on October 10, 1906, at Longford, Kansas. He received his primary education there, at Burlington, Kansas, and at Manhattan, Kansas. He received his secondary education at Manhattan, Kansas, where he was graduated from the public high school in June, 1924. His undergraduate college studies were undertaken for one year at Kansas State College, and were continued at the University of California at Los Angeles, where he received the Bachelor of Arts degree with honors in June, 1932. He began his graduate studies at the University of California at Los Angeles, and there he received the Master of Arts degree in June, 1934. His graduate studies were continued at the University of Chicago, where he was in residence for eleven quarters between October, 1934, and August, 1938. In 1936 he was married. Concurrently with his activities as a graduate student he was a teaching assistant in mathematics at the University of California at Los Angeles. In Chicago he taught two courses at the University College of the University of Chicago during the school year 1934-1935, and he was an evening division instructor in mathematics at the Armour Institute of Technology during the school years 1936-1938. His graduate courses were taken from Professors Hedrick, Whyburn, Garver, Barnett, and Ellis at the University of California at Los Angeles, and from Professors Bliss, Dickson, Lane, Lunn, Graves, Barnard, Reid, Hestenes, MacMillan and Zachariasen at the University of Chicago, to all of whom he is very grateful. He particularly desires to express his appreciation to Professors Bliss and Hestenes for their inspiring and watchful guidance in the development of this thesis.

